Electronic Version
Stylesheet Version v1.1.1

Description

[METHOD OF FABRICATING FLASH MEMORY]

BACKGROUND OF INVENTION

[0001] Field of the Invention

[0002] The invention relates in general to a method of fabricating a flash memory, and more particularly, to a method of a flash memory in which an overlap area between a floating gate and a control gate is increased.

[0003] Related Art of the Invention

[0004] Flash memory has been broadly applied in personal computer and electronic products due to the superior data retention characteristics.

[0005] The typical flash memory has a stack-gate structure, which comprises a tunneling oxide layer, a polysilicon floating gate used to store charges, a silicon oxide/silicon nitride/silicon oxide (ONO) dielectric layer, and a polysilicon control gate used to control the data access.

[0006] Normally, the larger the gate-coupling ratio (GCR) between the floating gate and the control gate, the lower the operation voltage required. Consequently, operation speed and efficiency are greatly enhanced. The method of increasing the gate-coupling ratio includes increasing the overlap area, or decreasing the thickness of the dielectric layer, or increasing the dielectric constant k of the dielectric layer between the floating gate and the control gate.

[0007] As mentioned above, increasing the overlap area between the floating gate and the control gate is advantageous to increasing the gate-coupling ratio. However, due to the continuous demand of higher integration, the area occupied by each memory cell has to be reduced. Therefore, how to fabricate a flash memory with a high gate-coupling rate within limited chip area has become an important task.

SUMMARY OF INVENTION

[0008] The present invention provides a method of fabricating a flash memory in which the overlap area between a floating gate and a control gate is increased, such that the coupling ratio thereof is increased.

[0009] The method of fabricating a flash memory provided by the present invention comprises the following steps. A tun-

neling dielectric layer, a conductive layer and a mask layer are sequentially formed on a substrate. The mask layer, the conductive layer and the tunneling dielectric layer are patterned to form longitudinally arranged strips on the substrate. Buried drain regions are then formed in the substrate between the strips. The strips are further patterned into floating gate structures, which thus comprise the patterned tunneling dielectric layer, the patterned conductive layer and the patterned mask layer. An insulation layer is formed on perimeters of the floating gate structures. The insulation layer has a top surface lower than a top surface of the patterned conductive layer of the floating gate structures, such that a part of the sidewalls of the conductive layer is exposed. The patterned mask layer is removed to expose the top surface of the patterned conductive layer, a gate dielectric layer is formed on the top surface and the exposed sidewalls of the patterned conductive layer, and a control gate is formed on the gate dielectric layer.

[0010] In the present invention, the height of the insulation layer formed on perimeters of the floating gates is reduced, such that the gate dielectric layer is formed on both the top surface and the sidewall of the floating gate. As a re-

sult, the overlap area between control gate formed on the gate dielectric layer and the floating gate is increased, and the gate-coupling ratio is increased.

Brief Description of Drawings

[0011] These, as well as other features of the present invention, will become more apparent upon reference to the following drawings.

[0012] Figures 1A to 1G are top views showing the fabrication process of a flash memory according to a preferred embodiment of the present invention.

[0013] Figures 2A to 2G are cross-sectional views along line I-I" as shown in Figures 1A to 1G, respectively.

Detailed Description

[0014] Figures 1A to 1G are top views showing the fabrication process of a flash memory according to a preferred embodiment of the present invention. Figures 2A to 2G are cross-sectional views along line I-I' as shown in Figures 1A to 1G, respectively. Referring to Figures 1A and 2A, a substrate 100 is provided. The substrate 100 includes a silicon substrate, for example. A tunneling dielectric layer 102, a conductive layer 104 and a mask layer 106 are sequentially formed on the substrate 100. The material of

the tunneling dielectric layer 102 includes silicon oxide, and the thickness thereof is about 50 angstroms to about 100 angstroms, for example.

[0015] The method for forming the tunneling dielectric layer 102 includes thermal oxidation or low-pressure chemical vapor deposition (LPCVD), for example. The material of the conductive layer 104 includes doped silicon formed by low-pressure chemical vapor deposition with silane as a gas source to deposit a polysilicon layer, followed by an dopant implantation process, for example. The operation of the deposition process is about 575°C to about 650°C, and the operation pressure thereof is about 0.3 torr to about 0.6 torr.

[0016] The material of the mask layer 106 includes silicon nitride or silicon oxide formed by low-pressure chemical vapor deposition using dichloro-silane and ammonia as reacting gas.

[0017] Referring to Figures 1B and 2B, a patterned photoresist layer 108 is formed on the mask layer 106. The mask layer 106, the conductive layer 104 and the tunneling dielectric layer 102 are etched using the patterned photoresist layer 108 as a mask to form a plurality of strips 200 longitudinally arranged on the substrate 100. The strips

200 comprise the patterned tunneling dielectric layer 102a, the patterned conductive layer 104a, and the patterned mask layer 106a. An ion implantation step is performed to form buried drain regions 110 in the substrate between the strips 200.

[0018] Referring to Figures 1C and 2C, the patterned photoresist layer 108 is removed. Another patterned photoresist layer (not shown) is formed on the strips, and the strips 200 are etched using the patterned photoresist layer as a mask to form the floating gate structures 300. The floating gate structures 300 comprise the patterned tunneling dielectric layer 102b, the patterned conductive layer 104b, and the patterned mask layer 106b. The patterned conductive layer 104b is then used as the floating gate.

[0019] Referring to Figures 1D and 2D, an insulation layer 112 is formed on the substrate 100 to cover the floating gate structures 300 and fill the spaces between the floating gate structures 300. The material for forming the insulation layer 112 is different from that of the mask layer 106b, including silicon oxide, silicon nitride or spin-on-glass, for example. The method for forming the insulation layer 112 includes high-density plasma chemical vapor deposition (HDP-CVD) or spin-coating, for example.

[0020] Referring to Figures 1E and 2E, the insulation layer 112 on the mask layer 106b is removed to expose the patterned mask layer 106b. The remaining insulation layer 112 located between the floating gate structures 300 is denoted by a reference numeral 112a. The method for partly removing the insulation layer 112 includes chemical mechanical polishing (CMP) or etch back, for example.

[0021] Referring to Figures 1F and 2F, a part of the remaining insulation layer 112a is removed to remain the insulation layer 112b with a top surface lower than a top surface of the conductive layer (floating gate) 104b, such that a part of the sidewall of the conductive layer 104b is exposed. The method for partly removing the insulation layer 112a includes etch back, for example.

[0022] Referring to Figures 1G and 2G, the mask layer 106b is removed to expose the top surface of the conductive layer 104b. The method for removing the mask layer 106b includes wet etching, for example. When the material of the mask layer 106b is silicon nitride, the etchant used in the wet etching step includes phosphoric acid.

[0023] A gate dielectric layer 114 is formed over the substrate to cover the top surface and the exposed sidewall of the conductive layer 104b. The material for forming the gate

dielectric layer 114 includes silicon oxide/silicon nitride/silicon oxide, for example. The method for forming the gate dielectric layer 114 includes forming an oxide layer by thermal oxidation, followed by low-pressure chemical vapor deposition to form a silicon nitride layer and another silicon oxide layer, for example. The gate dielectric layer 114 may also be made of other materials such as silicon oxide or silicon oxide/silicon nitride.

[0024] A conductive layer 120 is formed on the gate dielectric layer 114 as a control gate. The conductive layer 120 includes a polycide layer consisting of a doped polysilicon layer 116 and a silicide layer 118, for example. The method for forming the doped polysilicon layer 116 includes an in-situ doping step. The silicide layer includes a low-pressure chemical vapor deposition using metal fluoride and silane as the gas source, for example. The subsequent process for forming the flash memory is known in the art and is not further described.

[0025] Accordingly, as the height of the insulation surrounding the floating gate is reduced, such that the sidewall of the floating gate is partly exposed, resulting in a larger area for forming the gate dielectric layer and for overlapping with the control gate. Consequently, the gate-coupling

ratio is enhanced without increasing the cell area of the flash memory, such that the device integration is maintained.

[0026] Other embodiments of the invention will appear to those skilled in the art from consideration of the specification and practice of the invention disclosed herein. It is intended that the specification and examples be considered as exemplary only, with a true scope and spirit of the invention being indicated by the following claims.